DOCUMENT RESUME

ED 044 391 24 SP 004 448

AUTHOR	Alkin, Marvin C.
TITLE	Statewide Evaluation of Education of Education in New Mexico: Assessment and Recommendations.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	CSE-WP-11
BUREAU NO	BR-6-1646-03
PUB DATE	Oct 70
CONTRACT	OEC-4-6-061646-1909
NOTE	24p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.30
DESCRIPTORS	*Accreditation (Institutions), *Citizen Participation, *Evaluation, *Evaluation Criteria, Evaluation Methods, Evaluation Needs, *Public Schools, Research Needs

ABSTRACT

This report, presented to the Legislative Finance Committee of the State of New Mexico, assesses the present procedures for evaluating and accrediting the public schools and makes recommendations for changes. The major flaws found in the present procedures are the amount of man hours involved and the focus on processes rather than outcomes. Recommendations are made for a reorganization of the staff involved in evaluation to reduce the time required and to provide for more citizen participation. It is also recommended that evaluation emphasize outcomes wherever possible and that a research effort be undertaken to provide firm data on the relationship between certain educational processes and outcomes. An appendix contains selected questions and answers which followed presentation of the report. (RT)

ED044391

STATEWIDE EVALUATION OF EDUCATION IN NEW MEXICO: ASSESSMENT AND RECOMMENDATIONS*

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

MARVIN C. ALKIN

CSE Working Paper No. 11

October 1970

Center for the Study of Evaluation

UCLA Graduate School of Education

Los Angeles, California

*A report presented to the legislative Finance Committee of the State of New Mexico, Santa Fe, N. Mex., May 20, 1970

SP004448

In this report I will deal with five major topics: (1) I will present, for your consideration, an evaluation framework; (2) I will describe the New Mexico evaluation procedure; (3) I will discuss the New Mexico legislation prescribing evaluation; (4) I will talk about the evaluation functions that might be served in New Mexico; and (5) I will make some recommendations for your consideration.

Evaluation Framework

In order to establish a common frame of reference about evaluation and problems pertaining to evaluation, I would like to discuss with you some of my views about the nature of evaluation and the evaluative function. My colleagues and I at the Center for the Study of Evaluation at UCLA feel that past definitions of evaluation which have been accepted professionally have been somewhat inadequate. Evaluation is not simply tests and measurement, as it seems to be in the view of many. Tests and measurement are, however, components of the complex evaluation process. Evaluation, in the view of some, appears to focus exclusively upon the use of self-evaluation and expert judgement; and while these again are necessary components, they are not the total world of evaluation. Further, some have thought of evaluation as simply the specification of objectives, in which the function of the evaluator is to note whether or not the objectives have been achieved. While this is an important function, it is also not evaluation in and of itself.

What, then, is evaluation? We view evaluation as being distinct from research in that (a) the evaluator has in mind at all times the idea that he is serving a decision maker, and (b) evaluation itself has as its purpose the provision of information for decision making. Evaluation is the process

of deciding on the appropriateness of the decisions concerned, selecting information appropriate to those kinds of decisions, collecting and analyzing that information, and reporting that information in a form potentially useable by decision makers in making choices among alternatives.

With this viewpoint of the evaluation process in mind, we may then consider kinds of decisions involved in the instructional improvement process. One such decision is choosing among objectives. For example, relative to the objectives that might be considered important, which of the various functions served in a particular school district are deficient? More specifically, if there is money in the school district which we want to invest "where the wheel squeaks," we must know what constitutes a "squeak" and which "squeaks" will prove to be important. This stage of the evaluation process we call a needs assessment. Here we are concerned with examining the gap between specific goals and the existing situation, and then determining a procedure for selecting priorities.

A second kind of evaluation deals with decisions related to program selection. Given an objective area in which the system is deficient, the decision maker must determine which of the existing or alternative programs is most likely to be beneficial in achieving the stipulated objectives. Such choices are referred to as "program selection decisions," and are in large part based upon information provided in the corresponding program planning.

Once a program has been selected and has begun to be implemented, several kinds of evaluation, paralleling several key decisions, are required. One would be what we call implementation evaluation, which is concerned with the extent to which the program was indeed implemented in the manner specified. Some questions that might be asked at this point include: Does the original

description of the program fit the program as implemented? Have changes taken place which were not anticipated? (In which case we would not actually be evaluating the original program, but rather evaluating the program as modified by unanticipated changes.) Has the program, which was probably designed to fill the needs of a particular population, been implemented to the population for which it was intended? (The program might well have been correctly implemented, but to the wrong population, in which case its objectives would not have been achieved.)

Progress evaluation involves the provision of information to the decision maker regarding how the program is functioning during the process of its introduction. Here, the evaluator is concerned with providing information about the potential modification and improvement of the program. Thus, the evaluator during this stage is an interventionist, and not simply a researcher who is careful not to intercede. (From our point of view, children are too valuable to justify the detached viewpoint of the researcher at this point.) If the program is not functioning properly, the evaluator has the responsibility of providing that information to the decision maker. In so doing, it is true that the generalizability of the program, which the pure researcher would be anxious to maintain, may be destroyed. However, a malfunctioning program would have minimal generalizability in any case.

The fifth stage in the evaluation process is more conventional and tends to be very similar to a strict control research study. This stage we call outcome evaluation. The evaluator is no longer an interventionist and, in fact, it is imperative that he not be actively involved, since his role now is to aid the decision maker in considering the potential generalizability of the program--that is, the extent to which the program can be implemented in

other systems in order to achieve the purposes specified.

To summarize, the five stages of evaluation and their corresponding key decision areas are: needs assessment (problem selection); program planning (program selection); implementation evaluation (program operationalization); progress evaluation (program improvement); and outcome evaluation (program certification). For our purposes today I will focus on the first two--needs assessment and program planning--because these areas are the most relevant to your interests as legislators and to the purposes of this meeting today.

Since it is important for our further communication, let me provide you with definitions of three terms that I use frequently: inputs, processes (or process characteristics), and outputs.

There are three kinds of inputs to educational systems. Student inputs include the characteristics of the students that enter the system. Community inputs include the kinds of learning experiences that exist outside the school. Finally, financial inputs are given in terms of the dollar amount that the superintendent has to work with. All of these inputs are things over which the superintendent has little or no control.

Process characteristics are the interactions that occur between the various inputs in the process of producing the educational outcomes. The process characteristics, generally, are the kinds of things that are mentioned in the official minimum standards for New Mexico schools, such as student-teacher ratios, etc.

By educational outputs, we mean two things: (a) student outputs, or the cognitive and behavioral changes which take place in students after they are exposed to the program, and (b) non-student outputs, or the impact

of the program upon systems external to it, such as the home, community, other programs, etc.

The New Mexico Evaluation Procedure

The New Mexico evaluation procedure is an accreditation procedure modelled after that used by most educational accrediting agencies in the United States, including the North Central Association and the Western Association of Schools and Colleges. First among the key elements of this procedure are self-evaluation and the benefits which are supposed to accrue from self-evaluation. The New Mexico Department of Education provides to each of the school districts a self-evaluation guide, which consists of checklists, questions to be answered, ratios to be determined, etc.

The second major characteristic or key element of the New Mexico accreditation procedure is its focus on process variables or process characteristics, in terms of the definition above. As a result, both the procedure and the resulting report are very much oriented towards an examination of those process characteristics which are considered relevant and important. The self-evaluation guide is designed to appraise the educational system on the basis of the following kinds of information: (1) inputs, including school district and individual school data, community data, and finanacial data, and (2) process characteristics, including administrative support and procedures, staffing, program organization, and support services. Thus, the emphasis is primarily on inputs and processes.

The third major element is the use of on-site review teams. These vary in size according to the size of the district under review, the largest being the 42-member team sent to Albuquerque. The site review teams are composed primarily of staff of the New Mexico Department of Education.

The fourth major characteristic of the accreditation process is the site visitors' evaluation report, which is written by the site team members after spending a week or so with the staff in the local school district going over their self-evaluation guide (which has focused primarily on process variables).

This leads us to a very distressing aspect of the whole procedure. If there are 30 school districts per year being evaluated, with an average of about 20 members per site visit team, and if each visit entails about eight days' time (five days actual visitation in the school district, one day of preparation, two days writing the report), then you have 30 x 20 x 8, or roughly 4800 man days per year consumed in the assessment process. (I am sure that the State Superintendent or someone on his staff could give you a more precise estimate of this.) However, this does not include the considerable time spent by committees within the school districts in preparing the self-evaluation guides.

This might not be an inordinate amount of time if self-evaluation yielded the benefits that its advocates maintain it does. Unfortunately, the process variables on which it focuses are of dubious value. We really do not know, for example, how much difference a teacher ratio of 23 instead of 25, or of 38 instead of 40, makes. Nor do we have adequate research evidence as to whether ten square feet more in an art room will make the difference between the students having or not having artistic sensitivity. Thus, in my judgement, we really do not know the validity of most of these process characteristics; this is not a problem unique to New Mexico. As I noted earlier, the procedure implemented here is a variation of the most commonly used accreditation procedure in the country. (It must be admitted,

however, that a greater than average number of team members are employed in New Mexico.) The Department of Education instructional specialists with whom I have spoken have estimated that roughly 75% of their time is spent on the statewide evaluation process.

I would guess that such an amount of time away from home might cause family difficulties. But a possibly more vital consideration is that these individuals are instructional specialists--experts in this area--who ought to have the time available to work with school districts on the solution of their instructional problems. If we can think of the accreditation procedure as a combination of needs assessment and program planning, in terms of the stages that we discussed earlier, then the first kind of recommendation that I would make is that the procedure ought to focus on needs assessment in order to free the time of the instructional specialists. These specialists would use the outcomes of that assessment to aid school districts in solving their problems. The evaluation report, then, would mark the beginning stage of the work of the instructional specialists, rather than entailing up to 75% of their time.

New Mexico Legislation Related to Evaluation

I would like to comment now on the New Mexico legislation related to evaluation. In an attempt at understanding the nature of your evaluation program and its origin, I have been provided with and have read Chapter 180 of the Laws of 1969, Section 77-2-2W. This section defines one of the duties of the State Board of Education as, to "assess and evaluate for accreditation purposes at least one-third of all public schools each year through visits by department personnel to investiage the adequacy of pupil gain in standard required subject matter, adequacy of pupil activities,

functional feasibility of public school and school district organization, adequacy of staff preparation, and other matters bearing upon the education of qualified students." The first item, dealing with the adequacy of pupil gain in standard required subject matter, falls into the category of outputs, in terms of the three definitions mentioned earlier. The other items mentioned in the legislation relate to processes. Thus, the section specifies two kinds of things--outputs and processes.

The assessment, which has been in part delegated by the State Board to the Department of Education, is no easy task. As has been noted, the nature of the questions to be investigated are of two types. One deals with the educational outcomes of the schools of New Mexico (such as the adequacy of pupil gain in standard required subject matter). The second deals with the process characteristics of those educational programs which contribute to the achievement of the stipulated educational outcomes (such as the functional feasibility of public school and school district organization, and the adequacy of staff preparation). These are different kinds of evaluation questions and require information of different types. We will consider each of these issue separately.

1. The adequacy of educational outcomes: There is a necessity for determining the outcomes of the schools thoughout the state. But the appropriate outcome dimensions to be examined must be based upon the objectives generally considered to be relevant to education in New Mexico. This state government, as that of all states, has the legal responsibility for education. But, in New Mexico, you have provided a particularly high proportion of the total educational expenditure from state sources. Thus, it seems to me that in doing so, the legislature must express its concern

for a demonstration of fiscal and educational responsibility on the part of local school districts.

The evaluation reports of the New Mexico Department of Education state that a "generally accepted basic principle of evaluation (is that the) evaluation must be made in terms of the processes or values of the individual school." While this is as difficult to disagree with as "motherhood and country", it should not preclude the consideration of broader purposes or values relevant to education in New Mexico, generally. An evaluation procedure must be provided which demonstrates the extent to which the schools of New Mexico have achieved or are achieving at the minimum standard prescribed by the state along the educational objective dimensions considered to be relevant and appropriate for all children in New Mexico, irrespective of the district of their residence. Surely, learning to read at some minimal level cannot be considered as a purpose of education in New Mexico subject to determination strictly by local option. Moreover, neither objectives related to citizenship, nor other basic objectives which are part of the set of commonly accepted and desired objectives and goals of education can be so considered.

The priorities and objectives for education in New Mexico need to be established in clear operational terms with a view to reflecting the thoughts of all appropriate constituencies of the citizenry of New Mexico. By this statement I do not mean that the local school district should not have the option of establishing its own objectives. I do mean, however, that there is a common core of values, goals, and aspirations considered to be reasonable objectives for all children of New Mexico. These objectives ought to be specified and included as part of an evaluation procedure. This

is one area where I am in disagreement with current practice in New Mexico.

To illustrate this, let me take an example from the current statewide evaluation procedure. In each of the evaluations that I have seen, the target (or goal) for art education has been stated as follows:

> The art program benefits the entire student body. The program is to be thought of as more than courses and activities confined to an art workroom. The art program seeks to have all students know of the relationship of art to every important aspect of daily living. Art can help all students to develop manipulative and organizational skills in expressing ideas, feelings, and moods; to explore personal interests and aptitudes; to acquire a knowledge of man's visual art heritage for the purpose of building an understanding of our culture in relation to other times and places; to become involved in and respond to visual art experience; to make sound visual judgements suited to their maturity level; and to develop sensitive discrimination in the use of art at home, in school and in the community.

While I am individually aware of the important role that art education plays in the total program, I am nonetheless somewhat bewildered by this statement, on several counts.

(a) Since it apparently appears in all Department evaluation reports, one must wonder about the extent to which current evaluations truly are responsive to "purposes or values of the individual school"--that is, local determination.

(b) If the statement was written, as I understand it was, by the Department of Education specialist in art, can we assume that his views reflect the value placed upon art education by the citizens of New Mexico for all their students?

(c) I think you will agree that the statement as just read to you is certainly not operational in that one could test or measure the extent to

which the target (or goal) had been achieved.

(d) Apparently evaluation of individual schools is based on the perceptions of Department of Education staff members rather than on any specific quantitative data along clear-cut dimensions. However, most of the recommendations in these reports are addressed to suggested modifications in school procedures with only infrequent mention or analysis of the extent to which the individual school or district has achieved the "target." That is to say, most of the recommendations do not address themselves to the target, but instead to processes. This is not surprising, inasmuch as the targets are generally non-measurable.

2. Processes for achieving stated objectives: We previously noted the second major kind of evaluation activity specified in the legislation. This deals with an examination of the educational processes. The problem in this area is that there is generally inadequate evidence as to the appropriateness of various processes for achieving specific objectives. Thus, in an evaluation emphasizing processes, one is frequently faced with the dilemma of critiquing a school on its failure to utilize a specific process (for example, arts and crafts instructional equipment may be needed for the Albuquerque district); to expand or modify a process (such as lowering the pupil-teacher ratio, which is mentioned in most reports without any regard to the possibility of doing that by virtue of the funds that might or might not be available); or to eliminate the use of an instructional process. (I find very few examples of this latter recommendation.) All of this takes place without any specific empirical evidence as to which of these processes make a difference in the production of the desired educational goals or outcomes. This is not intended as a criticism of New Mexico or the New Mexico procedure, but rather as a

criticism of the state of educational research generally.

Functions of Evaluation in New Mexico

I would like to describe some of the possible purposes that evaluation in New Mexico might serve. After discussing four kinds of possible functions, we will turn to what might be done. In terms of the scheme of evaluation that I have presented, it seems to me that you are interested in focusing upon needs assessment evaluation in the statewide assessment. That is, you want to know how you are doing relative to what you expected to be doing. For this, it is necessary to focus on outcomes rather than on processes.

1. You might perform what I would call a statewide needs assessment. The first step in this procedure would be to establish a number of objectives considered to be important on a statewide basis, which would be used as the primary measure for evaluating education in New Mexico. The Legislature, as the decision maker, would be provided with information about how the schools of New Mexico are doing relative to the expecations of the Legislature and of the total citizenry of New Mexico. Moreover, considering the characteristics of the various communities of the state, not every school would be expected to achieve the same objectives at the same time. That is to say, you might establish differential norms for the schools of the State dependent upon certain socio-economic and educational factors.

2. A second function that might be possible is a "local needs assessment," which would focus upon objectives specified by individual schools, beyond those objectives established by the State. Ways of measuring these locally specified objectives would be established by the individual school districts, which would be judged not only on how well they achieved the statewide objectives, but also

on how well they achieved the locally determined objectives. However, I would say that such a local needs analysis would be difficult to establish because of the problems of measuring such objectives in every district.

3. Another important function that might be served by an evaluation program in New Mexico would be to check on the extent to which process characteristics known to be relevant, and which are mandated by the state or required by the Department of Education, are actually being satisfied. For example, one State established process characteristic is that the district superintendent should have the responsibility for the budget. From what I understand, there may be some very small schools districts in which this mandate is not satisfied. On the other hand, some of the State process characteristics are of such vital concern and are so generally well accepted that they tend to become objectives in and of themselves. Therefore, it seems to me that one would have to be exceedingly careful about the way in which such important processes were defined. One would also need to limit their number so as not to constrict the system with too many process characteristics, as is now the case. It would seem best to begin with a small, pre-stated list which would be expanded only gradually.

4. An additional function that I can see served on a statewide basis would be a research study on specific process characteristics that would parallel the annual evaluations. The purpose of this study would be to determine systematically the relevance of specific process characteristics to the achievement of stated objectives. Thus, in succeeding years, instead of working with those process characteristics you merely suppose are important, you will be able to work with those that have been proven as relatively important in the production of certain outcome dimensions.

To summarize, I see the four functions of evaluation in New Mexico as involving statewide needs assessment, local needs assessment, consideration of a small group of process characteristics legally required, and a parallel research study to determine their relevance to the achievement of stated objectives.

Recommendations

To implement an evaluation plan of the type described above, the following recommendations are made:

1. Statewide Goals Committee: I would recommend that you establish a statewide goals committee, by this summer [1970] if possible, to operate during this summer and fall. This committee would consist not only of members of the teaching profession, but of laymen as well. However, since professionals best know the means of implementing established goals, the procedures for measuring the extent to which those goals are being attained, and the instructional treatments that might be modified for increasing that level of attainment, it would be imperative that they also participate on the committee. The committee would hold public hearings throughout the state, with the intention of devising and specifying both the goals (and objectives) of education in New Mexico and the priorities of those goals. Priorities would have to be set because you might identify as many as 75 objectives, which would be a system so huge that in practice it would be unmeasurable. The statewide goals committee should act on a continuing basis so that the objectives can undergo modification as you work with the program and discover ways in which the objectives can be improved.

2. Executive responsibility: I would recommend that the person in charge of evaluation in the State Department of Education take responsibility

for acting as the executive officer of the goals committee, and that he also take responsibility for selecting and devising appropriate measures for evaluating the objectives specified by this committee.

3. Relationship to Educational PPBS: It is my understanding that you are now attempting to introduce PPBS into various State Departments. If the state should move into the PPB system in education, the objectives gathered by the statewide goals committee and the measures developed for these objectives would appropriately form the basis for the development of PPBS, because the PPB system demands the establishment of goals and objectives and the means to measure them. Moreover, the budget data derived from PPBS would enable cost-effectiveness analyses by objectives to take place.

4. Statewide testing of objectives: The establishment of a statewide testing program on a small number of objective dimensions would seem advisable as a means of determining how successfully the objectives are being attained. (Of course, some citizens will be concerned about why only those few objectives which you select are being considered. At least it would be a starting point, and the ensuing public discussion should provide valuable assistance for your future selection of additional objective dimensions.)

Test selection problems will arise in a number of ways. For example, you may need to develop tests for specific population groups--Spanish speaking citizens, Indians, etc. A second problem or area of potential concern relates to the inefficiency of considering a "total" testing program. You do not need to know how an individual student is doing; what is to be engaged in is program or system evaluation, not individual evaluation. To do this, you need to be involved in what we call matrix sampling--a sampling of tests given to a sampling of students. Working with aggregated data in this manner also

will reduce the costs of testing.

5. Legally required process characteristics: A list of process characteristics legislatively required or mandated by the Department of Education, known or assumed to be relevant, should be devised. This list would form a part of the evaluation procedure as outlined previously. The process characteristics, along with the outcome evaluation, would form the basis of the statewide assessment.

6. Associated Research Activities: I believe that you will want to learn more about various aspects of certain process characteristics. For this purpose you might designate a team of several persons to do research on those specific process characteristics in which you are interested. The data from this would not simply be filed away, but would be analyzed very carefully by the evaluation staff of the Department of Education along with the evaluation data already mentioned. This would provide further information about the processes important and relevant in New Mexico education.

7. Timing: Given test development and selection time, you would probably be able to do a needs assessment evaluation by next spring (1971) and report on it in the fall. Concurrently, the evaluation and research section of the Department of Education might begin working closely with school districts throughout the state in helping them to set their own objectives beyond those established by the state. Perhaps several years from now it will be possible for a local needs assessment component to be added to the statewide evaluation procedure, dependent upon the extent to which local objectives, as specified, have been measured. But this would be a very difficult task and it does not appear that it would be at all possible for this to be a part of your present evaluation scheme.

8. Other: The program would release specialists who are now spending

an enormous amount of time on the present assessment procedure, since the publication of the evaluation report would be only the beginning of their work. They would then be free to devote the greater portion of their time to an area where their skills are vitally needed--working intensively with districts that have problems in order to help them meet the required objectives.

While the present accreditation procedure used in New Mexico is probably performing as adequately as that used in most states, it can be greatly improved upon. A well thought out reorganization of the entire procedure should yield a more effective utilization of the total resources involved--financial, professional, and student--as well as increased qualitative output. Should the state of New Mexico make the decision to put into effect such a program, this more or less pioneer effort would set an example that I would hope to see widely emulated.

APPENDIX

Appendix

RESPONSES TO SELECTED QUESTIONS

Q: You mentioned that student-teacher ratios and the number of square feet per student in an art room have little correlation with the learning process. Will you explain this?

Alkin: I did not mean to imply that there is definitely little correlation. It is just that there is very little data on the relationships between characteristics like the number of square feet in an art classroom and the outcomes. But I am not saying that a smaller student-teacher ratio would not be correlated to educational outcomes. I am simply saying that the data are not clear enough. There is one area where the smaller student-teacher ratio does seem to be quite important, and that is on the primary level. But I am not advocating either a larger or a smaller ratio. This is just an example of the kind of thing for which we have very little data, which is nevertheless felt to be important.

Q: In other words, before we put into statute a 25 to 1 ratio, we ought to have better data than we have now. How do we get this better data?

Alkin: The ongoing research study which would parallel the evaluation program that I mentioned would be very necessary to provide the kinds of information that you gentlemen need as decision makers in order to know what kinds of things you want to put into the statutes and what kinds of prescriptive measures you want to place upon education. Meanwhile, in the absence of good data, you ought to be less restrictive rather than more restrictive and allow more flexibility in the system.

Q: Assuming that we were to accept some of the recommendations that you have made, how much more money are we talking about?

Alkin: I think that in the first year of the program, it could almost be done within the existing budget. A first step would be to increase the size of the research staff by several persons. It certainly would not seem reasonable to increase the size of the supervisors' staff, because 75% of their work load is to be shifted to another area. Without an increase in cost, it perhaps would be reasonable to draw research personnel partially from the supervisors' staff. I do not know what cost would be involved with the statewide committee but I do not expect that it would be particularly large. The statewide testing, however, would mean some increase in cost. But I strongly suspect this would overlap with and eliminate the need for many of the testing programs in the local school districts, so

you will eventually save money there, although I would see a necessity for some additional costs for the statewide testing in the first year.

Q: Would you consider it advisable to have private industry evaluate the schools? Would there be a private agency that could do this type of thing?

Alkin: I have seen private agencies that are less competent than most professional groups, and on the other hand, I suppose there are some private agencies that could do the job. I see no reason why the research staff of the State Department of Education would not be able to handle it. A large measure of the job deals with the establishment of goals which require working with the citizens of New Mexico. Thus, it would seem reasonable to have New Mexico people do the study rather than hiring outside agencies from elsewhere.

Q: Did I correctly understand you to say that you do not think we should have instructional specialists conducting field evaluations?

Alkin: That is correct. I think the evaluation program could be handled by expanding the research evaluation staff and placing a moratorium on the hiring of new specialists. The specialists now on the staff should concentrate on working with school districts in the area of instructional problems, rather than being involved in the evaluation process itself.

Q: Did you say that we need to set priorities on the money available?

Alkin: No, although it would be difficult to disagree with any statement about setting priorities on available money. But I believe that the statement I made dealt with the establishment of priorities on objectives. That is to say, your statewide goals committee undoubtedly will make up a list of objectives so large that it would not be physically possible to test every one of the dimensions on that list. Thus, there would be a necessity for setting priorities as to which of these objective dimensions should be considered in the first year's testing program, with the expectation that the testing program would undoubtedly increase in size as new objectives were introduced and became a part of the statewide system in the following years. But in the first year, because of the costs involved in selecting tests (and in some cases devising tests of your own), it would not be advisable to move too quickly into too many objective areas.

Q: Can you give us some examples of the sort of things that should be on this list?

Alkin: I would say that is the responsibility of the citizens of New Mexico to establish the important objectives for themselves. I do not think the Department of Education or the New Mexico Education Asso-

ciation should be in the position of telling the citizens of New Mexico what their expectations are for the goals of education.

Q: Then would you say that the Legislature should not be involved in telling the citizenry what the statewide needs are?

Alkin: One might consider the Legislature as being in a position to express the views of the citizenry as their duly elected representatives.

Q: In allocating money, do you feel that the Legislature should have strings tied to it to insure that the objectives are met?

Alkin: I think the Legislature has the responsibility for being aware of what is obtained for the financial resources that they provide, and I think that the evaluation system that we have been discussing today would provide that kind of information for the decision makers--the Legislature.

Q: You mentioned that the citizens of New Mexico should establish the objectives. How do you expect the populace to arrive at any feasible goals for education in the state?

Alkin: To say that the people of New Mexico should act as the final jurists in the selection of goals does not mean that they cannot call upon whatever expert advice they feel is necessary. But I would still maintain that the people of New Mexico are paying the cost and it is therefore their responsibility to indicate the kinds of outcomes that they expect from their educational program.

Q: Who do you think should appoint the statewide goals committee, and how many persons should be on the committee?

Alkin: I would say that the professional educators should be in the minority and that it should be primarily a citizens' committee. I do not know the legislative and administrative structure of this state well enough to know how such appointments would generally be made.

Q: As a matter of philosophy, should this committee be appointed by the State Board of Education working, presumably, under the State Department of Education, or should it be handled by our Legislative School Study Committee, which is an interim permanent committee of the Legislature?

Alkin: It seems to me that either might be appropriate. The Board of Education is entrusted with the responsibility for education in New Mexico and, therefore, establishing goals for education would not be an inappropriate activity for them to be engaged in. On the other hand, it would also not be inappropriate for the Legislative School Study Committee to want to examine the goals of education in New Mexico. Either alternative would seem reasonable, but local factors that are unknown to me might well indicate otherwide.

Q: I would like for you to give us your views on whether or not the summaries finally reached in our past research have been valid. Should we have dealt more with the outcomes?

Alkin: I have not seen these past research reports that have been done in New Mexico, but I would venture to say that, in general, process-oriented research which lacks a clear understanding of the objectives often is not very productive.

Q: Considering the practical realities of finances, how are we going to reach these goals once they are established?

Alkin: By having the objectives clearly in mind and used as a basis for the parallel research study discussed earlier, I believe that you will begin to get insights as to the kinds of procedures which are producing the desired outcomes in greatest quantity and most cost-effectively.

Q: From your review of the statute (Chapter 180, 77-2-2W), do you feel that it should be changed to make the evaluations more meaningful?

Alkin: It is possible that the evaluation procedure you are presently following may be incompatible with what is specified, since the statute talks about educational outcome dimensions and the present procedure, as you know, is concerned more with processes. The kind of evaluation procedure that I would propose would be concerned primarily with the outcomes, and the process characteristics would be subject to specific evaluation only as they were demonstrated to have relevance to the production of outcomes. As far as the language of the statute itself is concerned, the proposed evaluation could be conducted in a manner consistent with the statute. However, let us now consider the wording:

"...assess and evaluate for accreditation purposes at least one-third of all public schools each year <u>through visits</u> of department personnel..." Under the new procedure, visits would no longer be the prime modus operandi for collecting the data, which would be obtained through statewide testing, although there undoubtedly would be some visits taking place.

"...to investigate the adequacy of pupil gain <u>in standard required subject matter</u>..." This implies that the objectives are the same as the required subject matter. It also does not allow for differentiation between the relative importance of various objectives. I would personally prefer to have it read, "pupil gain <u>on objective dimensions considered relevant</u>."

There are a few other key phrases, that I consider troublesome. I am not quite sure what "adequacy of pupil activities" means. Does it mean the activities of students as they engage in the program? Does it mean their activities in extra-curricular events?

Moreover, unless the activity is related to some kind of outcome dimension, how is one to know whether or not it is adequate? Regarding "functional feasibility of public school and school district organizations": We might not know whether or not the organizational structure were adequate unless we examined it in light of the outcomes of the evaluation program and research study. The same comment would apply to "adequacy of staff preparation."

This statute represents an incomplete listing of process characteristics, without indicating the way in which one would be able to determine whether or not the process characteristics were appropriate.

Q: Is it necessary for us to duplicate the research of 49 other states, or can we lift the findings of similar research conducted in other states?

Alkin: There is no reason why you should not use as much of it as possible, but some of it will not be relevant, depending on various specific characteristics of the state of New Mexico and the sufficiency of data elsewhere.

Q: How large should the statewide goals committee be? Are you talking about a 50-man committee or a five-man committee?

Alkin: I would say more on the order of fifteen to twenty persons. It is my understanding that the Legislative School Study Committee has appointed an Ad Hoc Advisory Committee which has taken some first steps in this direction, although they do not yet have a complete list of goals which can be measured. There may be a possibility that this committee could be used as a model for the statewide goals committee. My main criticism of this Ad Hoc Advisory Committee would be that it is now so heavily manned with professionals with relatively little participation by lay persons.

Q: Are you familiar with the testing programs that we have had in New Mexico in the past? How would these compare with the sort of thing you are proposing? To follow up on the previous question about the tests that have been used in New Mexico: Some goals have been set by the Board of Education which are widely recognized--for example, arithmetic, language arts, and so forth. Now, if you have some tests of these specific things, would not this be an adequate way to test those goals or objectives?

Alkin: It depends on the extent to which the objectives have been satisfactorily stated in a manner that would allow the selection of the appropriate tests. If you give students a test simply because someone says, "I hear it's a good math test", you are accepting all of the objectives of that test, which are usually implicit, and you are not necessarily testing the students on the objectives which you previously established as relevant and important.

Q: Once you establish these goals, would there be some kind of follow-up after the students got out of school?

Alkin: If one of the objectives dealt with providing programs which will result in economically competent citizens (i.e., people that can hold a job), then that kind of data, rather than some kind of formalized test, might be the best means of gathering information relative to that objective. In other words, rather than a specific standardized test, there might be other information-gathering techniques more appropriate to some kinds of objectives.

Q: Where should the responsibility lie for establishing the written statement of broad educational philosophy--with the local administration, the local school, or with the State and the State Department of Education?

Alkin: I think there are some objectives that have to be considered as objectives for the total education program in New Mexico. Thus, a written statement of broad educational objectives should be established at the state level. But beyond that one has to constantly remember that every community is unique by virtue of the population it serves, the new kinds of local industries, and other factors. Therefore, there is a responsibility on the part of local school districts to establish aims and objectives which are peculiarly local in nature and which go beyond the objectives of the statewide program. The tests you have been using might or might not be appropriate. We would have no basis for knowing until we first specified the objectives.

Q: How long do you suppose the development of statewide objectives and program objectives at the local level will take?

Alkin: I think that the program objectives in the local districts is going to take very much longer than the development of objectives by the statewide committee on educational goals. But I believe that by using the material already developed by the Ad Hoc Committee of the Legislative School Study Committee and the work already done by the New Mexico Education Association, it would not be unreasonable to expect that some preliminary goal selection could be made and some tests administered in a statewide assessment by next spring.